

CARLA SÁ, RAYMOND J. G. M. FLORAX & PIET RIETVELD

(Received October 2005; revised September 2006)

**ABSTRACT** *This paper identifies pivotal factors behind individual decision making in the transition from high school to post-secondary education in the Netherlands. We apply a multinomial logit framework to individual data and accommodate two types of effects that have not received much attention in the literature. First, we analyse the impact of geographical accessibility of the higher education system. Second, we allow the individual observations to be correlated within schools, in effect accounting for localized social interactions. Our results confirm the paramount influence of the student's track record and talent. The results, however, also show that geographical proximity significantly increases the probability of high school leavers continuing their education at a university or professional college.*

## **L'accessibilité à l'enseignement supérieur importe-t-elle? Comportement du choix des diplômés de l'enseignement secondaire au Pays-Bas**

**RÉSUMÉ** *Cet article identifie les facteurs essentiels intrinsèques à la prise de décision individuelle dans la transition ayant lieu entre l'école secondaire et l'enseignement postsecondaire au Pays-Bas. Nous appliquons le cadre d'un logit multinomial à des données individuelles et nous adaptons deux types d'effets qui n'ont pas bénéficié d'une grande attention dans la documentation. Premièrement, nous analysons l'impact de l'accessibilité géographique du système de l'enseignement secondaire. Deuxièmement, nous permettons la corrélation des commentaires individuels au sein des écoles, qui représentent de fait des interactions sociales localisées. Nos résultats confirment l'influence primordiale des résultats obtenus de l'étudiant ainsi que de son talent. Les résultats montrent cependant également*

Carla Sá (to whom correspondence should be sent), NIPE, University of Minho, Portugal, and Tinbergen Institute, Gualtar, 4710-057 Braga, Portugal. Email: [cangelica@eeg.uminho.pt](mailto:cangelica@eeg.uminho.pt). Raymond J. G. M. Florax, Department of Agricultural Economics, Purdue University, West Lafayette, IN 47907-2056, USA, and Department of Spatial Economics, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Email: [rflorax@purdue.edu](mailto:rflorax@purdue.edu). Piet Rietveld, Department of Spatial Economics and Tinbergen Institute, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Email: [prietveld@feweb.vu.nl](mailto:prietveld@feweb.vu.nl). The first author gratefully acknowledges the financial support of the Portuguese Foundation for Science and Technology, FCT [SFRH/BD/5054/2001]. The authors wish to thank the Research Centre for Education and the Labour Market (ROA) for supplying the data used in this paper. Timo Huijgen and Maarten Wolbers of ROA have been extremely helpful in guiding the authors through the data acquisition process. The authors also benefited from the comments and suggestions of two anonymous reviewers of this journal, and discussions with participants of the 43rd Conference of the European Regional Science Association (Jyväskylä, Finland) and the ROA seminar (Maastricht, The Netherlands).

que la proximité géographique augmente considérablement la probabilité des sortants du secondaire qui poursuivent leurs études dans une université ou un collège professionnel.

## **¿Importa el acceso a la Educación Superior? La toma de decisiones en graduados de Educación Secundaria en los Países Bajos.**

**RESUMEN** Este trabajo identifica los factores claves en la toma de decisiones individuales en la transición del Instituto de Educación Secundaria a la educación Superior en los Países Bajos. Aplicamos un marco logarítmico polinomial a los datos individuales y analizamos dos tipos de efectos que no han sido estudiados detalladamente en este tipo de literatura. Primero, analizamos el impacto de la accesibilidad geográfica del sistema de Educación Superior. Segundo, permitimos que las observaciones sean correlacionadas dentro de las escuelas pudiendo explicar las interacciones sociales localizadas. Nuestros resultados confirman la significativa influencia de los antecedentes escolares de los estudiantes y sus habilidades. Los resultados, sin embargo, también revelan que la proximidad geográfica aumenta de gran manera la posibilidad de que los graduados de Instituto continúen su educación en una universidad o colegio profesional.

**KEYWORDS:** *High school graduates; higher education; social interaction; geographical accessibility*

**JEL CLASSIFICATION:** C25; I21; R10

### **1. Introduction**

The behaviour of high school leavers with respect to the choice between continuing education by entering post-secondary schools or entering the labour market has been analysed quite thoroughly. Initially, studies assumed the choice to be a simple binary decision between continuing schooling or entering the labour market (see, for instance, Kohn *et al.*, 1976). More recently, studies have considered broader sets of alternatives, including vocational education options and labour market alternatives, and have analysed the choice behaviour by means of multinomial models (see, for instance, Riphahn, 2002; Nguyen & Taylor, 2003; Giannelli & Monfardini, 2003).

The transitional behaviour of high school leavers is generally explained by individual characteristics, such as the individuals' capabilities, as well as the students' socio-economic background, usually measured by means of parental income, education, and occupation. Although most of the more recent studies include information on the spatial variability in labour market conditions, none of the studies seems to fully explore the spatial dimensions of the student's decision process as well as the potential relevance of localized social interactions.<sup>1</sup>

It is often assumed that characteristics of the regional surroundings of the parental household are the main source for variations in expected earnings and expected employability after schooling. Because some authors argue that it is quite unlikely that educational decisions are dominated by expectations related to the region where the student might possibly work after graduation (Hartog & Serrano, 2002), many studies introduce controls for spatial heterogeneity based on the regions where students live when making the decision whether or not to continue education. The spatial heterogeneity is generally related to labour market characteristics, although this seems to be an unnecessary restriction. Regional characteristics, such as the level of educational attainment, the intellectuality of the regional milieu, and the availability of local amenities may also be relevant (Sá *et al.*,

2004). In order to control for such aspects, some studies include regional dummies, usually derived from a rather crude division of the country into large heterogeneous entities such as North, South, East, and West (see Giannelli & Monfardini, 2003, in a study for Italy; and Nguyen & Taylor, 2003, in a study for the USA).

In most studies, individual students are the unit of analysis. However, individuals cannot be treated in isolation, and social interaction patterns should be accounted for (Manski, 2000; Brock & Durlauf, 2002). Although it is virtually impossible to identify one single social context that is most important for the student's choice behaviour after leaving high school, it is likely that the interaction in a cross-sectional data set of individual students can best be captured by assuming clustering among students attending the same high school. Students attending the same high school tend to be rather homogeneous in their socio-economic background, because they usually come from the same types of neighbourhoods. They also share various features of their everyday life, because they spend a significant part of the day attending classes together, they are being taught by the same teachers, and they are likely to spend some of their leisure time together. We therefore follow Moulton (1990) in maintaining that it is reasonable to expect that individuals who share an observational characteristic such as location (or high school) also share other unobservable characteristics, implying that disturbances are correlated.

We emphasize an additional spatial aspect in this paper, apart from accounting for spatial heterogeneity and localized social interaction as indicated above. Some previous studies have experimented with including distance to higher education institutions among the explanatory variables in the analysis of transitions of high school leavers (see, for instance, Kjellström & Regnér, 1998). However, none of the studies have considered the distance impediment in terms of a system-wide accessibility measure that includes all higher education institutions, eventually even distinguishing between professional colleges and universities. The latter is not only potentially relevant in explaining the choice behaviour of students but also provides important information for making informed policy decisions because, effectively, in most European countries the spatial distribution of higher education institutions is determined to a considerable extent by the national government (see Florax, 1992; Florax *et al.*, 2006; Sá *et al.*, 2004). In fact, the objective of policy makers is often to provide all potential students with the same higher education options and opportunities. In pursuing this objective, information on the spatial distribution of higher education participation is imperative.

We address the above issues using individual data on choices of high school leavers with a diploma, and combine these data with information on high schools and regional characteristics, for the Netherlands. Our main hypothesis is that individuals who live in closer proximity to a specific type of higher education institution (i.e. a professional college or a university) are more likely to continue studying after high school, and they are more likely to choose that type of institution, even after controlling for spatial heterogeneity and social interaction. We investigate the behaviour of Dutch high school leavers at the end of the last century (1998–2000) by means of a multinomial logit model distinguishing between university education, professional training, and no higher education as the main alternatives.

This paper continues as follows. Section 2 provides background information on the Dutch educational system. A state-of-the-art overview of the literature on choice behaviour of high school leavers is presented in Section 3. Sections 4 and 5

cover data issues, the empirical model, and the estimation results. Section 6 concludes.

## 2. The Dutch Educational System

Dutch students are admitted to secondary school after leaving primary education at an average age of 12, and they are essentially free to attend the high school of their choice, although parents are advised by the primary school as to the type of secondary education that best suits their children. Secondary education comprises schools providing practical training (PRO), pre-vocational secondary education (VMBO), general secondary education (HAVO), and pre-university education (VWO).<sup>2</sup> The longest track, VWO, takes 6 years and prepares students for university, although students with a VWO diploma do also have access to post-secondary professional colleges. Students from the HAVO track, instead, can only proceed to professional colleges, whereas those with a diploma for the VMBO track can go on to the MBO track, after which they can eventually attend higher professional education at professional colleges. Students have to choose one out of four programme profiles: science and technology, science and health, economics and society, or culture and society. The different profiles include both a set of courses common to all profiles (for instance, Dutch) and a series of courses specific to the profile.

The tertiary step of the educational system comprises the higher education sector. The Dutch higher education system is a dual system, with 13 universities (WO) and 50 vocational/professional colleges (HBO), which are almost all entirely publicly funded.<sup>3,4</sup> Every year, the national government determines minimum requirements regarding the secondary school diplomas that allow students to apply for post-secondary study programmes. Institutions can impose additional pre-requisites with respect to the profile and/or high school courses included in the diploma. In general, all students with a secondary school diploma have access to university education, that is, for most study programmes there are no supply constraints, although some exceptions apply.<sup>5</sup> The same is not entirely true for professional colleges, which tend to have somewhat greater autonomy in defining their admission criteria. Typically, they fix a broader range of entrance requirements, in particular related to skills, talent, or fitness for a profession.

Students are required to pay admission fees, which are equal across institutions and generally not very high. Quality differences between educational institutions are considered negligible in the Netherlands. Regular full-time students are eligible for student support provided by the national government, which is compatible with some part-time jobs. All students are eligible for a basic scholarship for the nominal duration of the higher education programme (4 or 5 years), the exact amount of which depends on whether the student lives with his or her parents. Depending on their parents' and their own income, students can also apply for additional funding, which is eventually supplied in the form of a supplementary grant or even a loan. Since 1990, all students have received a public transport pass, allowing free travel during workdays and discounted travel at weekends. Until the 1970s, a policy of geographical decentralization of the higher education system, resulting in the establishment of new universities, was implemented in the Netherlands, guided mainly by spatial equity concerns. As a result, the geographical accessibility of the university system is relatively high; there are about three universities per  $100 \times 100$  km grid cell. The same applies to post-secondary professional colleges,

the spatial distribution of which has traditionally been very even (see Florax *et al.*, 2006).

The above features of the Dutch higher education system, characterized as it is by inexpensive, spatially balanced and easy access to higher education, makes it less likely that price and supply considerations play a major role in the choice behaviour of students. In addition, given the high spatial density of institutions, it is questionable whether distance to such institutions is really an impediment.

### **3. Choice Behaviour of High School Leavers**

There is an extensive literature on the choice behaviour of high school leavers. We present a concise overview here, concentrating on aspects that have been identified as relevant to the decision whether or not to continue (post-secondary) education.

Human capital theory looks at education as an investment good. The decision to continue education depends on the anticipation that future returns for a post-secondary degree, over and above those for a high school diploma, outweigh the additional costs of extended schooling (including income forgone). Apart from being an investment decision, the demand for education can also be a current consumption choice (Kodde & Ritzen, 1988; Duchesne & Nonnemann, 1998). Students may attend college simply because they like the courses or the student lifestyle. Theories considering schooling as a consumption activity assume the demand for higher education to vary positively with student income and negatively with both direct (tuition) and indirect (opportunity) costs. Kodde & Ritzen (1984) integrate consumption and investment motives in a unique model, according to which students choose the optimal level of education, and current and future consumption, subject to time and budget constraints. The solution for the maximization problem suggests that the individual's demand for education is a function of direct and indirect costs, income and wage differentials.

The direct cost of attending a higher education institution has received considerable attention in empirical work. Direct costs include tuition, books and fees; expenditure for food and housing are not always considered because these exist in any case. The empirical literature shows that human capital investments are more likely when costs are lower (Bishop, 1977; Fuller *et al.*, 1982). The amount of financial aid, either in the form of grants or scholarships which cover at least part of the cost of college education, generally has a positive effect on the probability of enrolment (Fuller *et al.*, 1982; Catsiapis, 1987).

Household income is another important determinant of the decision to continue studying after secondary education. Most studies have found that the higher the household income, the higher the demand for post-secondary education as well as the propensity to be in school after the secondary level (see, for instance, Savoca, 1990; Duchesne & Nonnemann, 1998; Checchi, 2000; Hartog & Serrano, 2002). Educational attainment of parents and/or their occupational status are sometimes used either to proxy this income effect or to capture the independent positive influence it has on youngsters' decisions to attend higher education (e.g. Checchi, 2000; Hartog & Serrano, 2002; Nguyen & Taylor, 2003).

Average earning differentials between higher education graduates and high school graduates have been shown to be a good indicator of the relative labour market conditions. Empirical studies have found a positive effect of wage differentials between college and non-college occupations in local labour markets on the student's likelihood of attending post-secondary education (see, for example,

Duchesne & Nonnemann, 1998; Hartog & Serrano, 2002). The expectation of future unemployment, however, reduces the returns to education, and can therefore reduce the demand for higher education (Ordovensky, 1995; Riphahn, 2002). Current unemployment also plays a role in this decision process, with poor employment prospects retaining youngsters in school (Corman & Davidson, 1984; Savoca, 1990; Hartog & Serrano, 2002; Giannelli & Monfardini, 2003). Some recent studies also incorporate the effect of family, neighbourhood and ethnicity on individual human capital decisions, probably because knowledge of the behaviour of others reduces the risk and uncertainty involved in this type of decision (Borjas, 1995). Finally, human capital theory also predicts that, *ceteris paribus*, myopic people are less likely to go to college than forward-looking people, and that most college students are young (Ehrenberg & Smith, 2000). The degree of present-orientation is quite difficult to test, but age has been included in most empirical studies.

In addition to the consumption and human capital motive, participation in post-secondary education may also be related to higher education functioning as a screening or filtering device (see, for example, Kodde & Ritzen, 1984). While human capital theory suggests that education increases individual human capital, screening theory asserts that there is a selection effect at work. Participation in higher education is restricted to the more capable students, who also happen to be more productive. This is subsequently useful information for future employers, and higher education hence operates as a filter. Many studies use test scores as a proxy for individual talent and show that students with higher scores are more likely to attend post-secondary education (Fuller *et al.*, 1982; Venti & Wise, 1983; Catsiapis, 1987), in particular academic programmes (Ordovensky, 1995; Nguyen & Taylor, 2003).

Previous empirical studies have also found a series of other individual, family and school characteristics to be relevant. Gender seems to play a role in participation, but the results are not consistent across studies (Kodde, 1986; Kodde & Ritzen, 1988; Savoca, 1990; Ordovensky, 1995; Checchi, 2000). Race differentials are an important determinant of differences in college enrolment (Black & Sufi, 2002), as well as parental nationality and family structure (Nguyen & Taylor, 2003).

The type of secondary school that students attend may determine how likely the student is to enrol in higher education (Catsiapis, 1987; Kodde & Ritzen, 1988; Checchi, 2000; Nguyen & Taylor, 2003). The direction of this effect varies, however, between countries and with the structure of the educational system. The social status of the neighbourhood where the high school is located has a positive effect on youngsters' attendance of higher education institutions (Bishop, 1977). The educational track, the academic quality of the institution, and the plans of peers appear to have a positive effect as well (see, for instance, Fuller *et al.*, 1982; Savoca, 1990; Ordovensky, 1995). Finally, as far as spatial effects are concerned, the level of urbanization has been shown to play a role in determining choice behaviour (Riphahn, 2002; Giannelli & Monfardini, 2003; Nguyen & Taylor, 2003). Most studies have also found a negative distance effect (see, for instance, Fuller *et al.*, 1982; Ordovensky, 1995; Sá *et al.*, 2004), although there are exceptions as well (see, for instance, Kjellström & Regnér, 1998).

The above literature review shows that there is a vast series of potential determinants of the choice behaviour of high school students with respect to the decision to continue education (either at a professional college or a university) or choose another option, including entering the labour market. Typically, the

determinants are personal (including family background), school, and spatial characteristics. The overview also shows that it is important to take into account that decisions are not made in isolation, but rather within a network of social interactions. Finally, the overview shows that spatial effects and the impediments of distance should not be ignored. In the next section, we discuss how we include these aspects.

## 4. Empirical Framework

### 4.1. General Model

We model the choice behaviour of high school graduates on the basis of a utility maximization framework in a multinomial choice model. Let  $i$  represent the student,  $j$  stand for high school, and  $k$  indicate the geographical location of the student at the moment of graduation, and each student chooses between three different alternatives  $a$ , either: no higher education, go to a professional college, or enter a university programme ( $a = 1, 2$ , or  $3$ , respectively). The utility associated with alternative  $a$  is then given by:

$$U_{ijk}^{(a)} = V_{ijk}^{(a)} + \varepsilon_{ijk}^{(a)}, \quad (1)$$

where  $V_{ijk}^{(a)}$  is a linear predictor, and  $\varepsilon_{ijk}^{(a)}$  is a random term. Alternative  $f$  is selected over any other alternative  $g$  if  $U_{ijk}^{(f)} > U_{ijk}^{(g)}$ , for all  $g$ , or, equivalently,  $U_{ijk}^{(f)} - U_{ijk}^{(g)} = V_{ijk}^{(f)} - V_{ijk}^{(g)} + (\varepsilon_{ijk}^{(f)} - \varepsilon_{ijk}^{(g)}) > 0$ . If the error term  $\varepsilon_{ijk}^{(a)}$  has a Type I extreme value distribution (Gumbel), the differences  $(\varepsilon_{ijk}^{(f)} - \varepsilon_{ijk}^{(g)})$  have a logistic distribution, and it follows that the multinomial probability of response category  $f$  equals:

$$\pi_{ijk}^{(f)} = \frac{\exp(V_{ijk}^{(f)})}{\sum_{a=1}^A \exp(V_{ijk}^{(a)})}, \quad (2)$$

where  $\pi_{ijk}^{(f)}$  is the probability that  $f$  is chosen, and  $\sum_{a=1}^A \exp(V_{ijk}^{(a)}) = 1$ . In order to ensure identification of the model, we use alternative  $a = 1$  (no higher education) as the base category.

### 4.2. Data and Variables

The variables in  $V_{ijk}$  refer to personal characteristics of the individual indexed by  $i$ , characteristics of the high school that he or she attends indexed by  $j$ , at the spatial location of the moment of graduation indexed by  $k$ , the latter including accessibility to professional colleges or universities.

**4.2.1. Personal characteristics.** The data on student choices and other personal characteristics are derived from the RUBS survey (*Registratie Uitstroom en Bestemming van Schoolverlaters*) conducted by ROA (*Researchcentrum van Onderwijs en Arbeidsmarkt*) among students graduating from a pre-university high school (VWO; see Section 2), which is the only secondary education track that allows students to opt for professional or university higher education. The students are selected randomly in a stratified sampling process in which high schools are sampled first, and students in the subsequent second stage. We use survey data on 1998, 1999 and 2000 graduates, who responded to a questionnaire 18 months after graduation. The resulting sub-sample contains 3,263 observations.<sup>6</sup> The data set

contains the name and location of the high school, information about the respondent's main activity at the time of the survey (being a student, working, or out of the labour force), and, if applicable, the educational level attended. This information is used to create the choice variable (1 for no higher education, 2 for professional training, and 3 for university education), where the no higher education option is a mix of activities such as working, unemployment, out of the labour force, and non-tertiary education.

In line with previous studies, we derive information on personal characteristics from the sample, including gender, citizenship, parental citizenship, and age. We also obtain information on school performance by means of the mean grade point average (GPA). Finally, because high school programmes are organized in different 'streams' from which students may choose, we distinguish four profiles (science and technology, science and health, culture and society, and economics and society) by means of dummy variables.

We have data referring to three different cohorts of graduates, which we pool for estimation purposes. Since the distributions of variables tend to change over time, the identical distribution assumption may be not valid, although the independence assumption still holds. We therefore include cohort dummies in the econometric model to capture aggregate changes over time (see Wooldridge, 2002).

Descriptive statistics of the data according to the three dimensions of variation (personal, high school, space) are given in Table 1, and the distribution across the different choice categories is given in Table 2. Table 1 shows that the graduates are distributed evenly across the choice categories as far as gender, citizenship, age, and cohort are concerned. Table 2 shows that approximately 30% of the graduates choose a professional college, 65% continue education at a university, and 6% go elsewhere, mainly directly entering the labour market. The choice behaviour reflected in the sample information corresponds very well to the population. For 1999, for instance, the ministry reports that 26.2% of the VWO graduates chose professional training, 66.5% university education, and 7.2% went elsewhere (OCW, 2003, p. 49).<sup>7</sup>

*4.2.2. High school characteristics.* We include a limited number of high school characteristics in our analysis, and obtain these characteristics from yearly quality reports of each high school in the Netherlands, as conveyed in the evaluation of high schools by educational assessment authorities (*Inspectie Onderwijs*).

High schools in the Netherlands vary according to denomination. We distinguish public high schools, from private (non-religious) high schools, and private high schools with a religious denomination. We also include information on the size of the high school in terms of the total number of students, ranging from 426 to 3,020.

Tables 1 and 2 do not show large differences between the choice categories according to high school characteristics, except perhaps for school size. Specifically, in larger schools, graduates are more likely not to continue their education, and in smaller schools graduates are slightly more likely to go to university. There is a slight indication that graduates from private religious schools are more likely to choose a professional college.

*4.2.3. Spatial characteristics.* The data set does not contain information on family income and parental education. We therefore use areal data at the municipality



**Table 1.** Descriptive statistics by choice and personal, high school, and spatial characteristics, respectively (means, with standard deviations in parentheses for continuous variables)

Variables	Choice			Total
	No education	College	University	
Personal characteristics				
Male	0.3851	0.3368	0.4474	0.4119
Non-Dutch	0.0460	0.0200	0.0453	0.0380
Parents non-Dutch	0.0690	0.0316	0.0790	0.0647
Age	19.4023 (0.7126)	19.4421 (0.7085)	19.2941 (0.6581)	19.3429 (0.6793)
Grade point average (GPA)	6.5067 (0.5789)	6.5531 (0.4843)	6.9233 (0.6568)	6.7933 (0.6333)
Science and technology profile	0.2701	0.2379	0.4535	0.3809
Science and health profile	0.1322	0.1789	0.1388	0.1502
Culture and society profile	0.1494	0.1853	0.1267	0.1450
Economics and society profile	0.4483	0.3979	0.2810	0.3239
1998 graduate	0.3736	0.4526	0.4324	0.4352
1999 graduate	0.2874	0.2968	0.2768	0.2832
2000 graduate	0.3391	0.2505	0.2908	0.2816
High school characteristics				
Public school	0.2701	0.1895	0.2445	0.2298
Private school	0.0862	0.0684	0.0944	0.0864
Private religious school	0.6437	0.7421	0.6611	0.6837
School size	1,406.8450 (676.7423)	1,396.0270 (633.2048)	1,304.3450 (586.8493)	1,336.5030 (607.1513)
Spatial characteristics				
Income per capita	9.6879 (0.6762)	9.6343 (0.6184)	9.6550 (0.6737)	9.6508 (0.6582)
Population density	2,444.8620 (1,089.7470)	2,183.7280 (1,105.4400)	2,247.3660 (1,108.9640)	2,239.3700 (1,108.0300)
College accessibility	1.0295 (0.2944)	0.9677 (0.2927)	0.9724 (0.2992)	0.9741 (0.2973)
University accessibility	0.2899 (0.1455)	0.2580 (0.1402)	0.2597 (0.1250)	0.2608 (0.1309)
Groningen	0.0632	0.0747	0.0827	0.0794
Friesland	0.0172	0.0116	0.0131	0.0128
Overijssel	0.0575	0.0905	0.0519	0.0634
Gelderland	0.0057	0.0074	0.0089	0.0083
Utrecht	0.0287	0.0316	0.0243	0.0267
Noord-Holland	0.1724	0.1589	0.1725	0.1686
Zuid-Holland	0.5805	0.4642	0.4525	0.4628
Zeeland, N-Brabant	0.0287	0.1074	0.0935	0.0941
Limburg	0.0402	0.0442	0.0851	0.0708
Flevoland	0.0057	0.0095	0.0154	0.0132
No. of observations	174	950	2,139	3,263

level to capture these effects. The effect of the student's socio-economic background is included by using per capita income of the municipality where the high school is located.<sup>8</sup> The impact of cultural background and amenities is taken into account by including the level of urbanization of the municipality in which the high school is located, which is operationalized as population density. Both variables are obtained from CBS, the national statistics agency (CBS, 2003). Table 1 shows that there are no significant differences in per capita income across

**Table 2.** The distribution of graduates across choices, by personal, school, and spatial characteristics, respectively (in per cent), and the numbers of observation per category

	Choice			No. of observations
	No education (%)	College (%)	University (%)	
All observations	5.60	29.28	65.12	3,263
Personal characteristics				
Male	4.99	23.81	71.21	1,344
Female	5.58	32.83	61.59	1,919
Dutch	5.29	29.66	65.05	3,139
Non-Dutch	6.45	15.32	78.23	124
Dutch parents	5.31	30.14	64.55	3,052
Non-Dutch parents	5.69	14.22	80.09	211
Age $\leq 18$	6.48	21.30	72.22	108
Age 19	4.82	26.37	68.82	2,139
Age 20	6.36	35.05	58.58	833
Age $> 20$	6.01	38.80	55.19	183
GPA = 6	7.89	40.54	51.56	1,216
GPA = 7	4.06	26.02	69.92	1,576
GPA = 8	3.29	11.03	85.68	426
GPA = 9	0.00	0.00	100.00	45
Science and technology profile	3.78	18.18	78.04	1,243
Science and health profile	4.69	34.69	60.61	490
Culture and society profile	5.50	37.21	57.29	473
Economics and society profile	7.38	35.76	56.86	1,057
1998 graduate	4.58	30.28	65.14	1,420
1999 graduate	5.41	30.52	64.07	924
2000 graduate	6.42	25.90	67.68	919
High school characteristics				
Public school	6.27	24.00	69.73	750
Private school	5.32	23.05	71.63	282
Private religious school	5.02	31.60	63.38	2,231
Spatial characteristics				
Groningen	4.25	27.41	68.34	259
Friesland	7.14	26.19	66.67	42
Overijssel	4.83	41.55	53.62	207
Gelderland	3.70	25.93	70.37	27
Utrecht	5.75	34.48	59.77	87
Noord-Holland	5.45	27.45	67.09	550
Zuid-Holland	6.69	29.21	64.11	1,510
Zeeland, N-Brabant	1.63	33.22	65.15	307
Limburg	3.03	18.18	78.79	231
Flevoland	2.33	20.93	76.74	43

the different choice categories. The respective choice categories of no higher education, professional college, and university are inversely related to population density.

It is likely that other unobserved spatial characteristics play a role in the choice behaviour of graduates (for instance, regional production structure differences or cultural differences) and/or that the choices of graduates are spatially clustered. In order to avoid misspecification problems, we include dummies for the Dutch provinces as control variables.<sup>9</sup> In doing so, we account for spatial heterogeneity, which is altogether different from spatial dependence (or autocorrelation) that may

be relevant as well. One should note, however, that spatial heterogeneity and spatial dependence are generally difficult to disentangle in an empirical sense (Anselin, 2001; Florax & Nijkamp, 2005), and a corrective device for one of the misspecifications is likely to affect the seriousness of the other misspecification as well.

Finally, we incorporate two accessibility measures in our model, referring to spatial or geographical accessibility to professional colleges and universities, respectively. One should note that accessibility varies not only over space—which is obvious—but also over groups of individuals, because the eligibility to enter specific college and university programmes depends on the high school profile that the graduate adopts. Moreover, since we do not have exact geo-referenced information as to where the graduate lives, we use the distance between the graduate's high school and the respective colleges and universities that can be attended by students with a given profile. For ease of notation, we only use the subscript  $k$ , referring to space, instead of a more complex set of subscripts. Accessibility to universities is then defined as:

$$a(u)_k = \sum_{l=1}^{L_{i \in p}} \frac{1}{d_{jl}}, \quad (3)$$

where  $L_{i \in p}$  is the total number of universities offering study programmes for which student  $i$  who followed profile  $p$  in high school is eligible, and  $d_{jl}$  is the distance between the municipalities where the high school and the university are located, respectively. By analogy, we define the accessibility measure for professional colleges,  $a(c)_k$ . Accessibility measures are strictly positive,<sup>10</sup> and we assume that the higher the accessibility the greater the chance of choosing one of the educational alternatives.<sup>11</sup> Following utility theory for these kinds of models, the accessibility variables are included in logarithmic form (see, for instance, Rietveld & Bruinsma, 1998; Ortúzar & Willumsen, 2001).<sup>12</sup>

Table 1 shows that the geographical accessibility to professional colleges is substantially greater than to universities (0.97 vs 0.26), which is obvious given the total number of institutions of both types (approximately 50 vs 13).

#### 4.3. Utility Function and Econometric Aspects

The general formulation of the utility that each individual takes from each choice depends on individual, high school related, and spatial aspects, in the following way:

$$V_{ijk}^{(a)} = \alpha^{(a)} + \beta^{(a)'} x_i + \delta^{(a)'} \gamma_j + \gamma^{(a)'} z_k + \theta_1^{(a)} \log(a(u)_k) + \theta_2^{(a)} \log(a(c)_k), \quad (4)$$

where  $a = 1$  (no higher education), 2 (professional college), or 3 (university),  $x_i$  is a vector containing variables with personal characteristics,  $\gamma_j$  contains high school characteristics,  $z_k$  spatial characteristics, and  $a(u)_k$  and  $a(c)_k$  refer to university and college accessibility, respectively. For reasons of identification, the coefficients referring to choice 1 are all set to zero.

The utility associated with the choice of professional college education does not depend, however, on how accessible university alternatives are, and, *mutatis mutandis*, the same holds for the university alternative. This implies that we set  $\theta_1^{(2)} = 0$  and  $\theta_2^{(3)} = 0$  in order to obtain:

$$V_{ijk}^{(2)} = \alpha^{(2)} + \beta^{(2)'} x_i + \delta^{(2)'} \gamma_j + \gamma^{(2)'} z_k + \theta_2^{(2)} \log(a(c)_k) \quad (5)$$

and

$$V_{ijk}^{(3)} = \alpha^{(3)} + \beta^{(3)'} x_i + \delta^{(3)'} \gamma_j + \gamma^{(3)'} z_k + \theta_1^{(3)} \log(a(u)_k), \quad (6)$$

for the utilities associated with choosing professional colleges and universities, respectively. Moreover, owing to the inherent characteristics of the multinomial logit model, both accessibility measures appear in every choice probability (as utilities of each and every alternative appear in the denominator of equation (2)), allowing for possible substitution effects.

Before proceeding with the estimation of the above utility functions in a multinomial logit setting, we elucidate some salient econometric issues at stake. We have identified the potential of correlation among the individual observations because of network effects and/or spatial clustering. The former may be caused by social interaction in localized networks, and the latter by unobserved spatial characteristics such as regional labour market conditions. Lacking data prohibits the explicit incorporation of these phenomena in the model specification, and we therefore resort to accounting for the correlation by choosing an appropriate estimator. A straightforward way of accomplishing this is to use the Huber–White (‘sandwich’) estimator (see Wooldridge, 2002, section 13.8.2, for details). This is not possible, however, for spatial clustering, because a disjoint classification of spatially clustered individuals is at odds with the mere concept of spatial clustering.<sup>13</sup> We therefore use the modified Huber–White type of clustering to model localized social interactions. Specifically, we expect students attending the same high school to be more similar in their characteristics than individuals chosen randomly from the population. One can, of course, argue that social interaction and networks extend beyond the school’s boundaries—to the neighbourhood of residence, for instance. Since students attending the same high school tend to come from the same types of neighbourhoods and socio-economic contexts, we choose to define social interaction, and hence correlation among choices, by means of the high school attended by students.

As mentioned above, modelling spatial clustering (or dependence) in a multinomial logit model is somewhat cumbersome (see Fleming, 2004), and we therefore incorporate spatial effects by focusing on spatial heterogeneity through the inclusion of fixed effects for provinces. The fixed effects are intended to capture, among other things, regional labour market conditions. The fact that regional labour market conditions are likely to be correlated to per capita income and population density makes a fixed-effect specification preferable over a random-effects specification, because the correlation with the exogenous variables would cause the random-effects estimator to be biased.

Finally, it is important to discuss the potential sample selectivity problem of our data. It was noted in Section 2 that the tracking of students starts very early in the Dutch education system, because at the age of about 12 pupils have to choose one out of four possible tracks. The sub-sample used in this study considers only those individuals in the VWO track, which may suggest a problem of non-random sample selection in that some factors that have determined the track choice may also play a role in the higher education choice. This is the case if, for instance, an unobserved attribute like innate ability influences the choice in favour of the VWO track, and at the same time favours the following choice of university education.

Failing to account for sample selectivity implies that our estimates might suffer from sample selection bias. This is an obvious shortcoming of our study that we cannot correct for using a selection equation because we do not have data on students in all tracks and nor do we have a variable explaining the selection only. However, from the university standpoint VWO students constitute an important target group, as only these students can directly continue to university education. This suggests that this study does not lose its validity as a step towards understanding the role of space and the spatial distribution of higher education institutions to higher education choice. Our conclusions, however, cannot be extrapolated to students in tracks other than the VWO track.

In the next section, we present the estimation results for a multinomial logit model, with spatial fixed effects and modified Huber–White adjusted standard errors based on clustering of individuals attending the same high school, produced with the STATA 8.0 software.<sup>14</sup>

## 5. Estimation Results

The coefficient estimates of the multinomial logit model are difficult to interpret in that they refer to the effect of each variable on the log-odds ratio between the choice (professional college, or university) and the no higher education option. In order to avoid these difficulties, we compute marginal effects of each variable on each possible outcome (see Table 3). The reference case chosen in the computation of the marginal effects is a female Dutch student, with at least one Dutch parent, who chose the science and health profile and graduated in 1998 from a private school with a religious denomination, located in Zuid-Holland. All continuous variables are set to their sample means.

Table 3 shows that male students are more likely to go to university, and women to attend a professional college. The results also show that a student not being Dutch and/or having non-Dutch parents contributes to the odds of choosing the university alternative. The age of the student is negatively correlated with the university option, and positively with the professional college option, whereas the effect on choosing the no higher education option is not significantly different from zero.

The results for the different profiles seem to indicate that, as compared to the science and health profile, graduates with other profiles are more likely to choose the professional college option. Furthermore, over time, graduates have a tendency to turn away from academic training as compared to professional training or no continued training.<sup>15</sup> In accordance with previous studies, the student's high school performance and talents, as measured by the GPA, have unequivocally the largest marginal effect on the odds of choosing the university option. With respect to high school characteristics, there is a slight tendency for graduates from public as well as private high schools (although in the latter case the results are statistically not significant), as compared to graduates from private schools with a religious denomination, to choose the no higher education and the university education option as compared to professional colleges. It is unclear whether this is partly an artefact of the proportion of professional colleges with a religious denomination being higher than the proportion of universities with a religious base. The size of high school has no discernable effect.

The spatial characteristics of per capita income and level of urbanization are not significantly different from zero. This may be related, in part, to the relatively high

**Table 3.** Multinomial logit model, marginal effects

Variable	No education		College		University	
	Estimate	Robust SE	Estimate	Robust SE	Estimate	Robust SE
Personal characteristics						
Male	-0.0026	(0.0067)	-0.0718***	(0.0201)	0.0745***	(0.0201)
Non-Dutch	0.0140	(0.0191)	-0.0781*	(0.0429)	0.0641	(0.0459)
Parents non-Dutch	-0.0168**	(0.0085)	-0.1609***	(0.0257)	0.1777***	(0.0272)
Age	-0.0007	(0.0054)	0.0536***	(0.0146)	-0.0529***	(0.0155)
Log(GPA)	-0.2607***	(0.0578)	-1.1861***	(0.0963)	1.4467***	(0.1104)
Science and technology profile	-0.0041	(0.0124)	0.1403***	(0.0327)	-0.1362***	(0.0312)
Culture and society profile	0.0051	(0.0129)	0.1609***	(0.0398)	-0.1661***	(0.0370)
Economics and society profile	0.0191**	(0.0089)	0.1413***	(0.0293)	-0.1603***	(0.0273)
1999 graduate	0.0149	(0.0117)	-0.0074	(0.0196)	-0.0075	(0.0221)
2000 graduate	0.0166*	(0.0091)	-0.0378*	(0.0222)	0.0212	(0.0231)
High school characteristics						
Private school	0.0106	(0.0155)	-0.0439	(0.0376)	0.0333	(0.0371)
Public school	0.0275**	(0.0135)	-0.1036***	(0.0380)	0.0761**	(0.0366)
Log(school size)	0.0104	(0.0146)	0.0235	(0.0396)	-0.0339	(0.0396)
Spatial characteristics						
Log(income per capita)	-0.0270	(0.1109)	-0.3896	(0.2746)	0.4165	(0.2826)
Log(population density)	-0.0053	(0.0072)	0.0056	(0.0305)	-0.0003	(0.0305)
Log(college accessibility)	-0.0067*	(0.0035)	0.1103*	(0.0569)	-0.1037*	(0.0535)
Log(university accessibility)	-0.0089*	(0.0047)	-0.0535**	(0.0274)	0.0624**	(0.0320)
Groningen	-0.0404***	(0.0082)	0.0519	(0.0778)	-0.0115	(0.0776)
Friesland	-0.0257***	(0.0077)	0.0282	(0.0702)	-0.0025	(0.0699)
Overijssel	-0.0238***	(0.0076)	0.0315	(0.0533)	-0.0077	(0.0564)
Gelderland	-0.0284**	(0.0118)	0.0641	(0.0655)	-0.0357	(0.0609)
Utrecht	-0.0061	(0.0154)	0.0022	(0.0731)	0.0039	(0.0828)
Noord-Holland	-0.0140	(0.0131)	-0.0339	(0.0431)	0.0479	(0.0445)
Zeeland, N-Brabant	-0.0422***	(0.0067)	0.0434	(0.0547)	-0.0013	(0.0548)
Limburg	-0.0307***	(0.0073)	-0.0504	(0.0376)	0.0811**	(0.0357)
Flevoland	-0.0379***	(0.0052)	-0.0936**	(0.0447)	0.1315***	(0.0461)
No. of observations	3,263					
% Correctly predicted <sup>a</sup>	66.99					
Log pseudo-likelihood	-2,322.9451					
Pseudo-R <sup>2</sup>	0.10					

Notes: Significance at the 1%, 5% and 10% level is indicated by \*\*\*, \*\* and \*, respectively, with Huber–White adjusted standard errors based on correlation among graduates attending the same high school given in parentheses. The marginal effects are computed for a female Dutch student, with at least one Dutch parent, who has chosen the science and technology profile, and who graduated in 1998 from a private school with a religious denomination, located in Zuid-Holland. All continuous variables are set to their sample means.

<sup>a</sup> The percentage correctly predicted outcomes is computed as follows. For each observation: (i) we estimate the probability of each outcome; (ii) the outcome with the highest estimated probability is the predicted one; (iii) the outcome is correctly predicted if the predicted outcome is the observed outcome; and (iv) the percentage of outcomes correctly predicted is the total number of correctly predicted outcomes divided by the total number of observations in the sample.

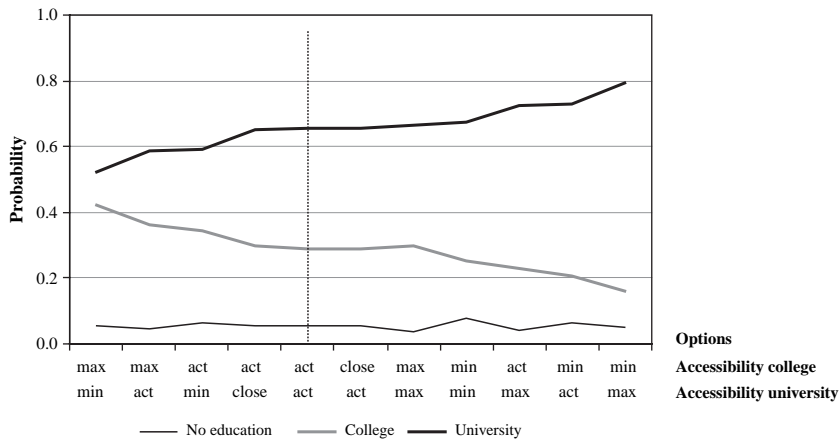
level of spatial aggregation that we use, given restricted data availability, and to correlation with the spatial fixed effects. With respect to the spatial fixed effects for the different provinces, again it should be noted that the marginal effects should be compared to the omitted category Zuid-Holland, which is one of the most dense and urbanized provinces in the Netherlands. Given the signs and significance of the

effects, it is actually likely that the dummy variables do pick up regional labour market differences as well as differences in other regional characteristics. Specifically, the marginal effects indicate that graduates living in rural areas (i.e. all provinces outside the highly urbanized Randstad, which is located in the provinces of Noord-Holland, Zuid-Holland, and Utrecht), in comparison to Zuid-Holland, have a tendency to prefer the education to the non-education option. Moreover, in all cases but two (Limburg, which has a very popular 'regional' university; and Flevoland, which is a polder very close to Amsterdam and Utrecht, both of which have large universities), they prefer to go to a professional college. Hence, there is a rather pronounced dichotomy between the Randstad area and the rest of the Netherlands.<sup>16</sup>

We end this discussion of the estimation results by looking more closely at the results related to geographical accessibility or distance deterrence. The coefficients of the accessibility variables show that geographical accessibility plays a significant role in determining the choices of youngsters in their transition from high school to post-secondary education or dropout from the educational system. Accessibility to professional colleges exerts a positive impact on decisions to continue with a professional education, while accessibility to university institutions has a positive influence on going to a university. A 1% increase in any of the accessibility variables hardly affects the probability of the non-higher education option (the probabilities decrease by only 0.0067 and 0.0089, respectively). If the accessibility to professional colleges increases by 1%, the probability of choosing that type of institution increases by 0.11, at the same time lowering the probability of choosing the university option by approximately the same amount (0.10). Similarly, although somewhat smaller in magnitude, a 1% increase in the spatial accessibility of universities increases the probability of choosing academic education by 0.06, with a concurrent decreasing effect of the choice for a professional college. The difference in magnitude of the effects between the two types of higher education illustrates that participation in professional training is more sensitive to changes in accessibility than university participation. It is important to note, however, that the number of universities is substantially smaller than the number of professional colleges (13 vs about 50). Hence, the number of new professional colleges required to bring about a 1% increase in accessibility is substantially higher than the number of new universities needed to accomplish a similar increase in university accessibility.

The effects of changes in accessibility are further illustrated by a series of simple simulations, the results of which are presented in Figure 1. The estimated average choice probabilities referring to the actually existing situation are 5.3% for the no higher education choice, and 29.1% and 65.6% for the choices of professional college and university, respectively.

The simulations in Figure 1 cover various situations. First, we give the reference case, with the actual (act) accessibility levels for both professional colleges and universities presented above (the vertical dashed line). Second, we compute the choice probabilities by fixing the accessibility to the minimum (min) or the maximum (max) level observed in the sample. And finally, we re-compute the accessibility level assuming that either one of the universities is closed down, or a professional college in one of the university towns is closed down. In the closedown case the choice probabilities shown in Figure 1 are the average of the different closedowns considered. The options in Figure 1 are grouped in ascending order of the probability of choosing to continue with a university education. The



**Figure 1.** Simulated choice probabilities for selected accessibility scenarios of professional colleges and universities.

*Note:* the abbreviations for the different options refer to accessibility fixed at the maximum value in the sample (max), at the minimum value in the sample (min), at the actual sample value (act), and at a re-computed value where one of the institutions is closed down (close).

figure clearly shows that changes in the accessibility of the higher education system have virtually no effect on the total number of high school graduates who decide to stay within or drop out of the higher education system. Hence, changes in the spatial accessibility of the higher education system have virtually no impact on the participation of youngsters in higher education as a whole. Changes in the accessibility of either compartment of the higher education system have a clear effect, however. The scenario where the accessibility of professional colleges is fixed at the sample maximum increases the likelihood of their choice, to the detriment of the university choice. *Mutatis mutandis*, similar effects hold for universities and for the scenario where the accessibility is fixed at the sample minimum. Finally, the figure shows that the effect of closing down either one university or one professional college (in a university town) has only a very small effect on national participation figures. At the local or regional level, the effect is, of course, notable, especially when the closure takes place in regions that are oriented strongly towards a particular institution. But, even then, the results indicate that the effects are rather limited, because most high school graduates opt to go to a different university rather than a nearby professional college.

## 6. Conclusions

Previous studies have documented that a wide range of personal, high school, and spatial factors determine the decisions of school leavers to continue their education or to drop out of the higher educational system. However, only limited attention has been paid to the potential relevance of localized social interactions and for the impact of space. The latter concerns both heterogeneity of the observed phenomenon over space as well as the potential distance deterrence effect that can be captured by accounting for spatial or geographical accessibility of the higher education system. We address these important issues in a case study of high school graduates in the Netherlands, during the period 1998–2000. We use a multinomial



logit model to investigate the choice behaviour of high school graduates 18 months after graduation, assuming that the school leavers have three options: university education, professional training, or no higher education. Localized social interaction is taken into account by allowing for observations of high school graduates attending the same high school to be correlated. Spatial effects are interpreted as spatial heterogeneity, and are taken into account by including fixed effects for areal units. The distance deterrence effect is incorporated through the inclusion of geographical accessibility indices for professional colleges and universities, respectively. Including variables reflecting both personal as well as high school characteristics mitigates omitted variable problems.

The empirical results confirm that past high school performance and talent of the high school graduates are strongly related to students' likelihood of going on to higher education. Another eye-catching result is that non-Dutch citizens and students with non-Dutch parents are more likely to choose to go on to university. Most importantly, we show that the choice behaviour of graduates has salient spatial dimensions.

Concurrently, population density has a relatively small effect on the likelihood of high school graduates going to a professional college. There is also a distinct dichotomy between the highly urbanized Randstad area and the rest of the Netherlands: graduates living in rural areas have a tendency to prefer the education to the non-education option, and they prefer to go to a professional college. The most outstanding spatial result is, however, that the geographical accessibility of the higher education system significantly contributes to high school graduates choosing to continue education. This effect is strongest for professional colleges: a 1% increase in the accessibility of professional colleges increases the odds of high school graduates choosing this option by 0.11. The corresponding effect for universities is about half the size (0.06).

We are aware of some shortcomings of our analysis. First, as we have pointed out in Section 4.3, the data we use potentially suffer from a sample selectivity bias problem, which we cannot fully correct with the available data. Therefore the results cannot be used to predict the behaviour of students in tracks other than the academic one. Second, we use population density and per capita income at the municipality level to proxy individual-level variables. The use of municipality averages results from lacking individual data. Although municipalities are in general not sufficiently homogeneous in terms of their socio-economic composition for these variables to serve as entirely adequate proxies, the use of these variables reduces the problem of omitted variable bias.

Our research can obviously be extended in various ways. It would be particularly useful to be able to incorporate information on the living arrangements of students (see, for instance, Martinez-Granado & Ruiz-Castillo, 2002), and to focus on the impact of supply constraints in the professional education tier of the higher education sector. Future research geared towards investigating the choice behaviour of prospective students based on precise geo-referenced individual data can contribute to using sophisticated spatial econometric techniques.

## Notes

1. Robertson & Symons (2003) are an exception as they consider the importance of localized social interactions on academic attainment.

2. The abbreviations are derived from the Dutch terms: practical training (PRO, *Praktijkonderwijs*), pre-vocational secondary education (VMBO, *Voorbereidend Middelbaar Beroepsonderwijs*), general secondary education (HAVO, *Hoger Algemeen Voorbereidend Onderwijs*), and pre-university education (VWO, *Voorbereidend Wetenschappelijk Onderwijs*).
3. Note that among the public higher education providers there are public and (almost completely) publicly funded non-profit private schools, which usually have a religious character. This distinction was the result of the movement of Protestants and Roman Catholics against the state.
4. Universities are referred to through the use of the abbreviation WO (*Wetenschappelijk Onderwijs*), and vocational or professional colleges are labelled HBO (*Hoger Beroepsonderwijs*). Over the last two decades, mergers have resulted in a sizeable reduction in the number of HBO institutions, falling from 350 in the mid-1980s to 56 in 2000, and subsequently to 50 in 2002 (OCW, 2003, pp. 74, 81).
5. For a limited number of profession-oriented programmes, such as medicine, dentistry, veterinary science, and information science, the national government fixes the number of students based on prospective demand in the labour market (*numerus clausus*).
6. The costs of the RUBS survey are partly born by the high schools, whose decision to participate in the survey is voluntary. The survey covers different schooling types, but we restrict the sample to graduates from VWO because those are the only students for which the three choices (i.e. no higher education, professional college, or university) are available. We use the 1999, 2000, and 2001 surveys, referring to the school year 1997/1998, 1998/1999, and 1999/2000, and include only those students who obtained a diploma and who supplied information on all relevant variables. See Potma & Kolk (2000), Potma (2002), and Huijgen (2002), for details on the surveys.
7. These figures are not perfectly comparable due to slightly different definitions and the moment at which the information was obtained (see OCW, 2003).
8. Municipal per capita income can be only a rough proxy of household socio-economic background, but unfortunately individual socio-economic data are not available for this sample. In the Netherlands, a municipality is a city or a group of small cities located very close to each other. In 2003, there were 489 municipalities in the Netherlands, with an average size of 85 km<sup>2</sup> and an average population of 33,114 inhabitants. One may also note, however, that in the context of the Dutch higher education system the relevance of household economic status is at least partly mitigated by the fact that all students are eligible for a basic scholarship, and it is only the supplementary grant that depends on family income.
9. We have experimented with different levels of spatial aggregation (for instance, the so-called NUTS I and II levels), but have found that a low level of aggregation, with many fixed effects as a result, explains away most of the variation. We therefore include 10 dummy variables for the 12 Dutch provinces because one of the provinces (Drenthe) is not represented in the data set, and the provinces of Zeeland and Noord-Brabant are pooled because of the low number of observations in the ‘no higher education’ category.
10. In order to avoid scale problems we define the intrazonal distance, which is relevant when a high school and a college or university location coincide in the same municipality, as:

$$d_i = ((\pi - 1)/\pi) \cdot \sqrt{s_i/\pi},$$

where  $s_i$  is the area of region  $i$  measured in square metres (see Rietveld & Bruinsma, 1998). The formula assumes that regions are circular, and all zones are used equally intensively. Although various alternative intrazonal measures are possible as well (see Sá *et al.*, 2004), the choice of one or the other does not have any serious bearing upon the results.

11. These are gravity-type measures, as they weight the number of opportunities of higher education of each type with the inverse of distance as impediment factor. We have also experimented with simpler accessibility measures such as the number of institutions of each type in the region and the distance to the nearest college/university. These measures are, however, highly negatively correlated with the accessibility measures, and the estimation results, therefore, barely change.
12. Rietveld & Bruinsma (1998, pp. 36–37) describe this type of accessibility measure in the context of the use of infrastructure services. The accessibility of a facility in a transport network is the expected value of the maximum utility of visiting that facility, which is assumed to depend on the mass of the facility, the travel costs of a trip to that facility, and a stochastic term. If the stochastic term is Weibull distributed, stochastic utility theory presumes accessibility measures that are typically of the form  $A = \log \Sigma_j \exp(\text{utility's deterministic part})$ , where  $A$  refers to accessibility, and  $j$  to destinations.
13. The Huber–White estimator requires the identification of observations that belong to clusters or groups of correlated observations (see, for example, Rogers, 1993; Williams, 2000, for details). This is perfectly feasible if one assumes network effects among individuals belonging to the same group or network. However, the nature of spatial dependence means that all observations, regardless of whether the observations refer to an individual or to an areal unit, typically belong to the same group. This can be seen as follows. In the spatial econometrics

- literature, spatial correlation is often modelled using contiguity or distance to determine spatial interaction. If, for instance, correlation is suspected among areal units that are contiguous, and unit *a* is correlated to unit *b*, and unit *b* to unit *c*, then unit *a* and *c* end up belonging to the same group of correlated observations regardless of whether or not they are contiguous, simply because they have a mutual link to unit *b*.
14. Using the STATA `mlogtest` command, we tested for the Independence of Irrelevant Alternatives (IIA) assumption, which is implicit in multinomial logit models. The null hypothesis was not rejected by the Hausman test, and there is therefore no statistical evidence against IIA.
  15. A possible reason for this is that the originally rather generous system of student financial assistance has been subject to several changes, leading, in effect, to lower grants, increased parental contributions, and a check on students' progress (Boezerooy, 2003).
  16. This is consistent with earlier findings on the basis of a gravity model using aggregate areal data (see Sá *et al.*, 2004).

## References

- Anselin, L. (2001) Spatial econometrics, in: B. Baltagi (ed.) *Companion to Theoretical Econometrics*, pp. 310–330, Oxford, Blackwell.
- Bishop, J. (1977) The effect of public policies on the demand for higher education, *Journal of Human Resources*, 12(3), 285–307.
- Black, S. E. & Sufi, A. (2002) *Who Goes to College? Differential Enrollment by Race and Family Background*, Working Paper 9310, National Bureau of Economic Research, Cambridge, MA.
- Boezerooy, P. (2003) *Higher Education in the Netherlands*, Country Report, Centre for Higher Education Policy Studies (CHEPS), Enschede.
- Borjas, G. J. (1995) Ethnicity, neighborhoods, and human capital, *American Economic Review*, 85(3), 365–390.
- Brock, W. A. & Durlauf, S. N. (2002) A multinomial-choice model of neighborhood effects, *American Economic Review Papers and Proceedings*, 92(2), 298–303.
- Catsiapis, G. (1987) A model of educational investment decisions, *Review of Economics and Statistics*, 69(1), 33–41.
- CBS (Centraal Bureau voor de Statistiek) (2003) *StatLine*, Statistics Netherlands. Internet site: <http://statline.cbs.nl>
- Checchi, D. (2000) University education in Italy, *International Journal of Manpower*, 21(3–4), 177–205.
- Corman, H. & Davidson, P. K. (1984) Economic aspects of post-secondary schooling decisions, *Economics of Education Review*, 3(2), 131–139.
- Duchesne, I. & Nonnemann, W. (1998) The demand for higher education in Belgium, *Economics of Education Review*, 17(2), 211–218.
- Ehrenberg, R. G. & Smith, R. S. (2000) *Modern Labor Economics: Theory and Public Policy*, 7th edn, Reading, MA, Addison-Wesley.
- Fleming, M. M. (2004) Techniques for estimating spatially dependent discrete choice models, in: L. Anselin, R. J. G. M. Florax & S. J. Rey (eds) *Advances in Spatial Econometrics: Methodology, Tools and Applications*, pp. 145–168, Heidelberg, Springer.
- Florax, R. J. G. M. (1992) *The University: a Regional Booster? Economic Impacts of Academic Knowledge Infrastructure*, Aldershot, Ashgate.
- Florax, R. J. G. M., Hall, P., Titheridge, H. & Wikhall, M. (2006) A comparative analysis of the geography of student recruitment and labor market entry, in: G. Törnqvist & S. Sörlin (eds) *The Wealth of Knowledge: Universities and the New Economy*, forthcoming.
- Florax, R. J. G. M. & Nijkamp, P. (2005) Misspecification in linear spatial regression models, in: K. Kempf-Leonard (ed.) *Encyclopedia of Social Measurement*, pp. 695–707, San Diego, Academic Press.
- Fuller, W. C., Manski, C. F. & Wise, D. A. (1982) New evidence on the economic determinants of post-secondary schooling choices, *Journal of Human Resources*, 17(4), 477–498.
- Giannelli, G. C. & Monfardini, C. (2003) Joint decisions on household membership and human capital accumulation of youths. The role of expected earnings and local markets, *Journal of Population Economics*, 16(2), 265–285.
- Hartog, J. & Serrano, L. D. (2002) *Earnings Risk and Demand for Higher Education*, Discussion Paper 122/3, Tinbergen Institute, Amsterdam.
- Huijgen, T. (2002) *Methodiek schoolverlatersinformatiesysteem 2001*, ROA–W–2002/3, Researchcentrum voor Onderwijs en Arbeidsmarkt, University of Maastricht.
- Kjellström, C. & Regnér, H. (1998) *Does Distance to a University Affect Enrolment Decisions? Evidence from Data on Three Cohorts of Swedes*, Working Paper 8/98, Institute for Social Research, University of Stockholm.
- Kodde, D. A. (1986) Uncertainty and the demand for education, *Review of Economics and Statistics*, 68(3), 460–467.
- Kodde, D. A. & Ritzen, J. M. M. (1984) Integrating consumption and investment motives in a neoclassical model of demand for education, *Kyklos*, 37(4), 598–608.

- Kodde, D. A. & Ritzen, J. M. M. (1988) Direct and indirect effects of parental education level on the demand for higher education, *Journal of Human Resources*, 23(3), 356–371.
- Kohn, M. G., Manski, C. F. & Mundel, D. S. (1976) An empirical investigation of factors which influence college-going behavior, *Annals of Economic and Social Measurement*, 5(4), 391–419.
- Manski, C. F. (2000) Economic analysis of social interactions, *Journal of Economic Perspectives*, 14(3), 115–136.
- Martinez-Granado, M. & Ruiz-Castillo, J. (2002) The decisions of Spanish youth: a cross-section study, *Journal of Population Economics*, 15(2), 305–330.
- Moulton, B. R. (1990) An illustration of a pitfall in estimating the effects of aggregate variables on micro units, *Review of Economics and Statistics*, 72(2), 334–338.
- Nguyen, A. N. & Taylor, J. (2003) Post-high school choices: new evidence from a multinomial logit model, *Journal of Population Economics*, 16(2), 287–306.
- OCW (Ministerie van Onderwijs, Cultuur en Wetenschappen) (2003) *Facts and Figures 2003*. Internet site: <http://www.minocw.nl/english/figures2003/pdf/figures2003.pdf>
- Ordozensky, J. F. (1995) Effects of institutional attributes on enrolment choice: implications for post-secondary vocational education, *Economics of Education Review*, 14(4), 335–350.
- Ortúzar, J. D. & Willumsen, L. G. (2001) *Modelling Transport*, 3rd edn, Wiley, Chichester.
- Potma, L. (2002) *Methodiek schoolverlatersinformatiesysteem 2000*, ROA-W-2002/2, Researchcentrum voor Onderwijs en Arbeidsmarkt, University of Maastricht.
- Potma, L. & Kolk, P. (2000) *Methodiek schoolverlatersinformatiesysteem 1999*, ROA-W-2000/9, Researchcentrum voor Onderwijs en Arbeidsmarkt, University of Maastricht.
- Rietveld, P. & Bruinsma, F. (1998) *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy*, Heidelberg, Springer.
- Riphahn, R. T. (2002) Residential location and youth unemployment: the economic geography of school-to-work transitions, *Journal of Population Economics*, 15(1), 115–135.
- Robertson, D. & Symons, J. (2003) Do peer groups matter? Peer group versus schooling effects on academic attainment, *Economica*, 70(1), 31–53.
- Rogers, W. H. (1993) Regression standard errors in clustered samples, *Stata Technical Bulletin*, 13, 19–23.
- Sá, C., Florax, R. J. G. M. & Rietveld, P. (2004) Determinants of the regional demand for higher education in the Netherlands: a gravity model approach, *Regional Studies*, 38(4), 373–390.
- Savoca, E. (1990) Another look at the demand for higher education: measuring the price sensitivity of the decision to apply to college, *Economics of Education Review*, 9(2), 123–134.
- Venti, S. F. & Wise, D. A. (1983) Individual attributes and self-selection of higher education: college attendance versus college completion, *Journal of Public Economics*, 21(1), 1–32.
- Williams, R. L. (2000) A note on robust variance estimation for cluster-correlated data, *Biometrics*, 56(2), 645–646.
- Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA, MIT Press.